DOCUMENT RESUME

ED 168 441 HE 011 163

AUTHOR Eiszler, Charles F.; Stancato, Frank
TITLE The Connotative Meaning of College Grades.
PUB DATE Mar 79
NOTE 26p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Grades (Scholastic); Higher Education; *Semantic Differential; *Student Attitudes; *Student Evaluation; *Student Reaction; *Undergraduate Students

ABSTRACT

A semantic differential technique was used to determine the connotative meaning of specific grades to college students. The 161 upper division undergraduates rated each of three specific grades on 20 semantic differential items. Two descriptive questions were answered: (1) What are the basic dimensions of the connotative meanings of grades? (2) How does the continuum of connotative meanings (as measured by "semantic distances" between adjacent grades) compare to the quantitative continuum that grades are assumed to represent? Factor analysis identified four factors: evaluation, realism, complexity, and salience. Comparison of the continuum of connotative meaning and the quantitative continuum revealed discontinuities. Findings are discussed in relation to the general theoretical assumption that grades have meaning to the extent that they initiate in the student an internal process which mediates the intended value judgment of the teacher regarding the performance of the student as a learner. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED168441

THE CONNOTATIVE MEANING OF COLLEGE GRADES

Charles F. Eiszler and Frank Stancato

Central Michigan University

March, 1979

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Charles Eiszler

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HE 011 163

In recent years traditional grading systems have come under considerable fire at all levels from elementary to the university and professional schools. Student attitudes and feelings about grades have been a major concern. Many students feel victimized by traditional grading systems (Poole, 1975). Still others feel that grades are unrelated to how hard they work in a course or the degree of real understanding that they have about the subject matter (Hardy, 1974). Suggestions for resolving the grading controversy have ranged from the elimination of grades to the adoption of alternative grading systems.

A number of research studies have been interested in determining the effect of nontraditional grading practices on the student's attitudes or opinions about the course. Wegener (1977) found no relationship between the active participation of the student in the grading process and the student's attitude about the course. Correspondingly, findings reported by Ball (1973) Yarber (1974) and Newcomb and Warmbrod (1974) show no relationship between student contracting and the attitudes students held about the course. In addition, Burton (1977) found contract grading to have a negative effect on the attitudes of students toward almost every aspect of a course.

In response to student efforts over the years many colleges have introduced some form of a limited pass-fail system. Frankel (1974) Chase (1973) and Hardy (1974) concluded that in general students display a favorable attitude toward the pass-

fail system of evaluation. Moreover, in some college communities the three point grading system of honors, pass-fail was more acceptable than the five point ABCDF scale (Goldstein, 1971). On the other side Boring (1975) and Gatta (1976) report strong feelings of students against the pass-fail system.

The idea of criterion-referenced testing has been rekindled by the recent emphasis on behavioral objectives, competency-based education and the development of programmed materials. Williams (1975) found criterion referenced grading systems to have little or no effect on the attitudes students held about the grading system or the course as a whole. On the other hand results reported by Hambleton and Murray (1977) support the belief that both faculty and students are in favor of a criterion-referenced system.

Research on the student attitudinal response to alternative grading systems leads to no definitive conclusions. These studies have, for the most part, been conducted within a larger context in which a traditional grading system continued in effect. If they tell us nothing else, perhaps these studies imply that we have too little understanding of that larger context. Perhaps student attitudinal or affective response to alternative grades and grading systems is unpredictable, because the nature of student response to the traditional system is poorly understood.

During the period of criticism and innovation, the traditional approach to grading has not been without its advocates.

Ebel (1975) examined the 22 most common criticisms of grades. He found some of these mutually inconsistent with each other, many unsupported by experience or research, others either inaccurate or irrelevant. More recently, McKeachie (1976) outlined a rationale for the traditional grading system which hinges on the assumption that grades and grading systems are primarily forms of communication. He identified three implications of this position:

1. A teacher giving a grade is usually communicating with more than one individual.
2. What a teacher communicates will depend, in part, on the meaning of the grade to the person reading it.
3. Teachers cannot unilaterally control or change the meaning of a grade.

McKeachie's defense of the traditional approach is simply that it is traditional and, therefore, provides a common background of information for the communication process. But McKeachie also suggests an important and unanswered question about the traditional grading system: What do specific grades mean to those who read them?

The current study was an attempt to determine what the specific grades used as a part of one university's traditional grading system mean to students at that university. For the purpose of this study we have relied on the work of Osgood (1957)

and his associates to provide both a theoretical and operational definition of meaning. Theoretically, we have assumed that grades have meaning to the extent that they initiate in the student an internal process which mediates an intended value judgment of the teacher regarding the performance of the student as a learner. Operationally, we have assumed that the semantic differential technique developed by Osgood is adequate to measure the connotative meaning of grades, that is, to assess the qualities of the mediation process within the student which are evoked by a grade.

In the study we have restricted ourselves to an examination of the connotative meanings of specific grades used in a traditional grading system and limited our investigation to the meanings evoked in upper division college students enrolled in a teacher education program. Two specific purposes of the study were identified:

1. To describe the basic components of students connotative meanings for specific grades.
2. To compare the continuum of connotative meanings with the quantitative continuum implied in the grading process.

METHOD

Sample

One hundred and sixty-one students taking a course in

educational psychology as a part of their elementary teacher education program were surveyed. Seventy-eight percent of the students participating in the survey were females and seventy-five percent were in their final semester as undergraduate students. The median accumulated grade point average for these students was approximately 3.2 on a four point scale, or in letter grades, between a B and a B+. These students were enrolled in five of seven sections of the course taught during the Winter and Spring Sessions of 1978.

Instruments

To assess the affective meaning of grades, twelve semantic differential scales were created, one for each of the twelve grade symbols used in the university grading system. Each scale was composed of a statement identifying the grade to be rated (Getting a grade of "C" in a college level course, for example.) and twenty pairs of polar adjectives: positive-negative, foolish-wise, good-bad, unimportant-important, successful-unsuccessful, weak-strong, severe-lenient, soft-hard, active-passive, slow-fast, exciting-calming, confusing-clear, predictable-unpredictable, simple-complicated, understandable-mysterious, impossible-possible, responsible-irresponsible, difficult-easy, fun-work, false-genuine. The twelve scales were randomly assigned to one of four forms: Form A (D, C, B-), Form B (B, C-, A), Form C (D+, B+, D-) and Form D (A-, E, C+).

Data Collection

Students responded to the survey during the final fifteen minutes of a regularly scheduled class period. Alternate forms of the survey were randomly distributed in each section. Consequently each student rated each of three specific grades on each of twenty semantic differential items.

Data Analysis

Student ratings of specific grades produced 483 sets of ratings on the 20 item set of polar adjectives. Correlations among the ratings on each item across all 483 sets were factor analyzed using the principle factoring with iteration method of the SPSS program (Nie, Hull, Jenkins, Steinbrenner, and Bent, 1975). A VARIMAX rotation of the principle factor matrix was used to determine the correlations of items with underlying factors and was the basis of our interpretations of the meaningful dimensions of student affective response to grades. Item-factor correlations greater than .40 were considered significant.

Semantic distances between adjacent grade evaluations were calculated using items selected to represent the dimensions identified in the factor analysis. The affective continuum created by these distances was scaled for comparison with the quantitative continuum which the grades are taken to represent in a quantitative or measurement sense.

RESULTS

The study was designed to provide descriptive data on two questions: 1) What are the basic dimensions of the connotative meaning given specific grades by students? 2) How does the continuum of affective or connotative meanings of grades compare with the denotative or quantitative continuum which grades are assumed to represent?

Dimensions of Affective Response to Grades

In the principle factor analysis, four factors were considered significant (eigenvalues greater than 1). These factors accounted for 65.8 percent of the variation in semantic differential ratings of grades. Rotation of the principle factor matrix resulted in factors which accounted for the following percentages of common variance: 60.2, 21.3, 12.1, and 6.4

Table 1 shows the semantic differential items which were significantly correlated with each factor and the item-factor correlation. Factor 1 correlates highly with items which reflect the evaluative nature of grades: bad-good (.884), negative-positive (.863), unsuccessful-successful (.848), irresponsible-responsible (.772), and foolish-wise (.648). The factor was named "evaluation" and the five items with the highest correlations with the factor were selected for use in subsequent analyses. Factor 2 correlates most highly with two items which seem to reflect the extent to which students perceive particular grades as realistic or possible for them:

impossible-possible (.651) and difficult-easy (.645). This factor was termed "realism" and the two items with highest correlations with the factor were used to represent the factor in subsequent analyses.

Insert Table 1 about here

Factor 3 is correlated with items which relate to perceptions of the complexity or understandability of grades. One item, simple-complicated (.784), was selected to represent this "complexity" dimension. A fourth factor, also represented by a single item, involved the "salience" of a grade: unimportant-important (.451).

Thus the factor analytic procedures which we used produced a pattern of interrelations among the 20 semantic differential items which suggested the existence of four major components to student meanings for specific grades. Students view grades evaluatively; some grades are better, more positive, and more indicative of success than others. Grades are also felt to have a quality of realism (or unrealism); some grades are less likely, less possible or less easy to attain than others. Student reactions to grades contain two other qualities: a sense that some grades are less understandable or more complicated in their meaning than others and a sense that some grades are more important than others.

In terms of the variance associated with each factor, the evaluation dimension is clearly the most important aspect of student response to grades. The evaluative dimension of grades accounted for almost three times the variance of the realism dimension, five times the variance of the complexity factor, and eight times the variance of the salience factor.

In selecting items for additional analysis, the items with the highest correlations with their respective factors were selected. An attempt was made to include items representing a particular factor in approximately the same proportion as that factor's proportion of the explained variance.

Table 2 shows the mean factor score for each grade on each factor. Data in this table show that the higher the grade the more positive the evaluative response, the lower the grade the less positive the affective response. On the realism dimension, a B and a B+ are seen as the most realistic grades, with both higher and lower grades rated as less realistic. In this regard, the semantic differential ratings of grades are consistent with the actual student experience as reflected in cumulative grade point averages. Generally, grades in the B range are also viewed as less complex or more understandable than either higher or lower ranking grades. With respect to salience, grades in the C range are viewed as the least salient and grades higher or lower are more important.

Insert Table 2 about here

In summary, the results of the current study identify four components involved in student affective responses to specific grades: evaluation, realism, complexity, and salience. These factors relate to the grade continuum in different and uniquely meaningful ways.

Comparing the Connotative and Quantitative Continua for Grades

To allow for more precise comparison of the affective meanings of grades in relation to the quantitative meanings involved in computing grade point averages and in the other ways in which grades are considered measures of student achievement, the semantic distances between the ratings for specific adjacent grades were calculated using a generalized distance formula (Osgood, 1957, p. 91). Table 3 shows the distances for each affective component, the total affective distance, and the adjusted or scaled distances for each pair of adjacent grades. The data show that the greatest discontinuity between the affective continuum and the quantitative continuum occurs for the distance between the grades of B- and C+. This affective distance is more than three times the quantitative difference. Generally, a B- is perceived to be closer to an A and a C+ is viewed as closer to an E than their positions on the quantitative continuum suggest. This result is more clearly presented in Figure 1 which portrays the distances graphically. In fact, this figure shows that generally every grade below a B- was perceived as closer in affective meaning to extreme failure than would be expected were the quantitative values

taken literally, and were psychological meaning consistent with mathematical meaning. Conversely, every grade at or above a B- was viewed as closer in affective meaning to extreme success than would be warranted by a literal interpretation of the grades as measurements

---

Insert Table 3 about here

Insert Figure 1 about here

In other words, there appear to be serious discontinuities between the continuum of affective responses to grades and the assumed measurement or quantitative continuum which they represent. The discontinuity suggests that students find grades of C+ and less as less psychologically satisfying than their quantitative values might suggest to instructors awarding these grades. Grades of B- or higher seem to be perceived as more psychologically satisfying than might be expected.

CONCLUSIONS AND IMPLICATIONS

For upper division undergraduate students enrolled in a elementary teacher education program, the affective response to specific grades, as measured by the semantic differential technique, includes four factors: evaluation, realism, com-

plexity and salience. In the research which culminated in The Measurement of Meaning (Osgood, 1957), studies of the connotative meaning of a wide variety of concepts, involving many respondent groups were conducted. In virtually all of these studies, an evaluative factor accounting for half to three-quarters of the extracted variance, was discovered. The data of the current study are consistent with the overwhelming body of semantic differential research, including the original studies by Osgood and his associates. The meaning of specific grades includes an "attitudinal" factor which reflects the sense in which the grade is viewed as a reward or a punishment.

Two factors identified in Osgood's original research and in many subsequent studies were not derived from the analysis of the data in the current study. These are a "potency factor" reflecting power and an "activity factor" a dynamic quality. Instead, three other factors, realism, complexity, and salience were extracted.

It should be noted that both the particular set of adjective pairs and the specific nature of the sample of respondents limits the generalizability of these results. Studies using more or different polar adjectives might be expected to yield a different outcome, but it seems likely the differences would be in the number and nature of the factors other than the evaluative factor. Research using a broader sample of college students, that is students following different curricular choices as well as students at different levels of completion of the educational program would also be helpful in extending the validity the conclusions of this study. Nevertheless, it seems

clear that the specific meanings of grades can be effectively assessed by the semantic differential technique and that to a very great extent the connotative meanings of grades involves a reward-punishment dimension.

For those who are upper division undergraduate students, there appears to be an important discontinuity between the connotative meaning of various grades and the meaning implied by the quantitative values associated with the grade. We believe that the data of the current study show that grades of C+ and less (except for that which denotes failure) are perceived by students to be more punitive than their rank or position in the system of grades would imply. Conversely grades of B- or higher (with the exception of an 'A') communicate more reward than is necessarily intended. Certainly the distance on what we have called the affective continuum between a grade of a B- and a grade of a C+ is far different than a teacher may recognize when he tries to decide between the two for labeling the achievement of a particular student.

Thus the results of the current study can be seen as consistent with the previously cited research of Boring (1975) and Gatta (1976) indicating a negative student reaction to pass-fail grading. If the 'pass' quality of the pass-fail system is seen as roughly equivalent to a grade in the C range, the distinction between the two, in an affective sense, reduces to a choice between two relatively aversive possibilities. Our research shows that even a C+ is perceived as closer to failure (an 'E') than to extreme success (an 'A').

Our research convinces us that grade inflation exists in a psychological sense. The student whose achievement in a class earns him a grade of 'C' probably feels less positive and less successful than his instructor intended in assigning the grade. Interestingly, the grades our students were most likely to get, those in the B range or higher, are the grades not devalued in their meaning. This raises the possibility that the acquisition of affective meanings for grades is a function of a context effect. Students define success primarily in terms of the modal levels of performance of their peers and associate positive affective meaning with grades assigned to their peers functioning at that level. One implication of such a model is that the affective meaning of a grade in one area of study or discipline may be different than the affective meaning of the same grade in another field.

Hewitt and Jacobs (1978) asked college students to estimate the grade-point average of fictitious students in eight different fields after reading a one paragraph description which included information on the student's sex, major field, and verbal and mathematical ability. Results of this study revealed that students perceptions of grading practices include the notion that different standards prevail in different major fields. In addition, student perceptions were generally consistent with the realities of the institution in which the study was conducted. Hewitt and Jacobs discuss the meaning of their data in relation to an "adaptation level hypothesis" as an explanation for differential grading standards in college fields of study. The basic theoretical assumption is that individual student performance is judged against the background of the abili-

ties of all other students in the class. If students' abilities in one field of study are different than another, grading standards will be different. Actual grade distributions may not vary. Hewitt and Jacobs suggest that student awareness of the difference in standards plays a role in the selection of a major field of study.

Our data suggests that the level of affect associated with a grade may depend, in part, on student judgments which can also be understood in terms of an adaption level hypothesis. Students judge grades attained by students at and above the modal level of their peers to connote success and such grades mediate the positive feelings associated with success. Conversely grades below those attained by the modal level of ones peers are viewed as varying degrees of failure and mediate negative feelings associated with the sense of failure and lack of success.

Taken together with the Hewitt and Jacobs research, our study suggests the following paradox. Students may select major fields of study, in part, because of awareness that standards for grading are different and that it's easier to attain higher grade-point averages in some fields than in others. Once in these fields, however, they evaluate their own success in reference to the other students in that field. Students who have chosen a field with lower standards and higher overall grade-point averages will have to attain higher grades to achieve the same degree of positive affect and sense of personal satisfaction that could have been attained with lower grades if they had chosen a field with higher standards.

A related concern involves the problems of an individual teacher whose standards differ from the prevailing standards in the field of study. New faculty or temporary instructors, particularly if they attempt master-based or criterion-referenced systems of evaluation and grading, may employ standards which differ from the prevailing standards. In addition, some faculty may consciously adopt more rigorous standards. Students whose grades are higher than the modal grade for the class, but lower than the modal grade for the field or area of study, may be expected to mediate less positive feelings than the instructor intends.

The limitations involved in the current study require that we be cautious attempting to draw implications for practice. Grades have connotative or affective meanings and these meanings can be described using a semantic differential technique. In addition, there seems to be a discontinuity between the affective meaning of specific grades and the quantitative meaning of the grades. The affective meaning of grades seems to be a function of the prevailing grading practices and standards and the data of the current study are consistent with a hypothesis that grade inflation has occured and has devalued the meaning so some grades, i.e., those in the C- range for example.

What other factors influence the affective meanings of grades? What instructional and evaluation practices will insure continuity between the affective and other meanings of grades? These are interesting and important issues but beyond the scope of the results reported here.

REFERENCES

Ball, Linda V., Student Contracting for Achievement Grades in Ninth Grade General Mathematics, (Ph.D. dissertation, The University of Connecticut, 1973).

Boring, Phyllis Z., Does standard grading encourage excessive competitiveness. Change, 1975, 44-46.

Burton, Gene E., Effects of contract grading on student evaluation of teachers. Educational Horizons, 1977, 55, 128-133.

Chase, Clinton, et.al., Pass-Fail: How It Is Working, (U.S. Educational Resources Information Center) ERIC Document ED 076161, 1973.

Ebel, R.L., Shall we get rid of grades? Measurement in Education, 1974, 5, 1-5.

Frankel, Edward, Student Faculty Evaluation Of A Three Point Grading System In Graduate Education Courses, (U.S. Educational Resources Information Center) ERIC Document ED 091999, 1974.

Gatta, Louis J., An analysis of the pass-fail grading system as compared to the conventional grading system in high school chemistry. Journal of Research In Science Teaching, 1976, 10, 3-12.

Goldstein, Kenneth, Attitudes toward ABCDE and honors pass-fail grading systems. The Journal of Educational Resources, 1971, 65, 99-100.

Hambleton, Ronald K. and Murray, John, A Comparative study of faculty and student attitudes toward a variety of college grading purposes and practices. (Paper presented at The Annual Meeting of The National Council On Measurement In Education - New York, New York, April 5-7, 1977.)

Hardy, Clifford A., Pass-fail as a system of evaluation. Improving College and University Teaching, 1974, 22, 242-243.

Hewitt, B.N. and Jacobs, R., Student perceptions of grading practices in different major fields. Journal of Educational Measurement, 1978, 15, 213-218.

McKeachie, W.J., College grades: A rationale and mild defense. AAUP Bulletin, 1976, 62, 320-322.

Newcomb, Lawrence H. and Warmbrod, Robert J., The Effect of Contract Grading On Student Performance, (U.S. Educational Resources Information Center) ERIC Document ED 093967, 1974.

References (Cont'd.)

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., and Bent, D., Statistical Package for the Social Sciences, Second Edition. New York: McGraw Hill, 1975.

Osgood, C.E., Succi, G.J., and Tannenbaum, P.H., The Measurement of Meaning. Urbana, Illinois: University of Illinois Press, 1957.

Poole, Richard, Student victimization and the formulation of test construction criteria. (Paper presented at the National Council On Measurement In Education - Washing-on, D.C., March 31 - April 2, 1975.)

Wegener, Elaine H., Use Of Student-Composed Tests and Their Effects On The Attitudes and Task Performance Of University Students. (U.S. Educational Resources Information Center) ERIC Document ED 129754, 1976.

Williams, Reed G., et.al., Differential effects of two grading systems on student performances. Journal Of Educational Psychology, 1975, 67, 253-258.

Yarber, William L., A comparison of the relationship of the grade contract and traditional grading methods to changes in knowledge and attitude. Journal of School Health, 1974, 44, 395-398.

Table 2. Mean Evaluation, Realism, Complexity and Salience Ratings for Specific Grades

Grades	Evaluation	Realism	Complexity	Salience
A	6.51[a]	4.74	4.35	6.33
A-	6.23	4.59	5.15	5.51
B+	5.78	5.11	3.68	5.71
B	5.68	5.16	4.03	5.40
B-	5.24	4.91	3.76	5.80
C+	3.65	4.77	4.28	4.36
C	3.18	4.29	4.22	5.12
C-	2.57	4.52	4.38	4.65
D+	1.80	3.57	4.90	5.71
D	1.81	3.11	5.66	5.63
D-	1.82	3.28	5.12	5.68
E	1.70	4.22	4.87	5.33

[a]On a 7-point semantic different scale.

Table 1. Item Factor Correlations

	Items	Factors			
		1	2	3	4
1.	Positive/Negative	.863			
2.	Foolish/Wise	-.684	.524		
3.	Good/Bad	.884			
4.	Unimportant/Important				.451
5.	Successfu/Unsuccessful	.848			
6.	Weak/Strong	-.628	.532		
7.	Severe/Lenient		.581		
8.	Soft/Hard			.605	
9.	Active/Passive	.696			
10.	Slow/Fast	-.536			
11.	Exciting/Calming	.546			
12.	Confusing/Clear	-.614			
13.	Predictable/Unpredictable			.456	
14.	Simple/Complicated			.784	
15.	Understandable/Mysterious	.520		.505	
16.	Impossible/Possible		.651		
17.	Responsible/Irresponsible	.772			
18.	Difficult/Easy		.645		
19.	Fun/Work			.588	
20.	False/Genuine		.531		

Table 3. Connotative Meaning Distances Between Adjacent Grade Values

	Adjacent Grades										
Connotative Dimensions	A/A-	A-/B+	B+/B	B/B-	B-/C+	C+/C	C/C-	C-/D+	D+/D	D/D-	D-/E
Evaluation (5)	.46	1.10	.16	1.39	12.65	1.30	2.21	3.56	.15	.12	.15
Realism (2)	.05	.83	.00	.18	.21	.50	.46	1.81	1.25	.80	2.83
Complexity (1)	.64	2.16	.12	.07	.27	.00	.03	.27	.58	.29	.06
Salience (1)	.67	.04	.10	.16	2.07	.58	.22	1.12	.01	.00	.12
Total	1.83	4.14	.38	1.81	15.20	2.38	2.92	6.76	1.98	1.21	3.16
Connotative Scale	.18	.40	.04	.17	1.46	.23	.28	.65	.19	.12	.30
Quantitative Scale	.30	.40	.30	.30	.40	.30	.30	.40	.30	.30	.70

NOTE: For comparison to the corresponding quantitative scale values, raw score distances were converted to appropriate proportions of a scale of 4.0. The revised distances are identified as the Connotative Scale.

.·(·.23

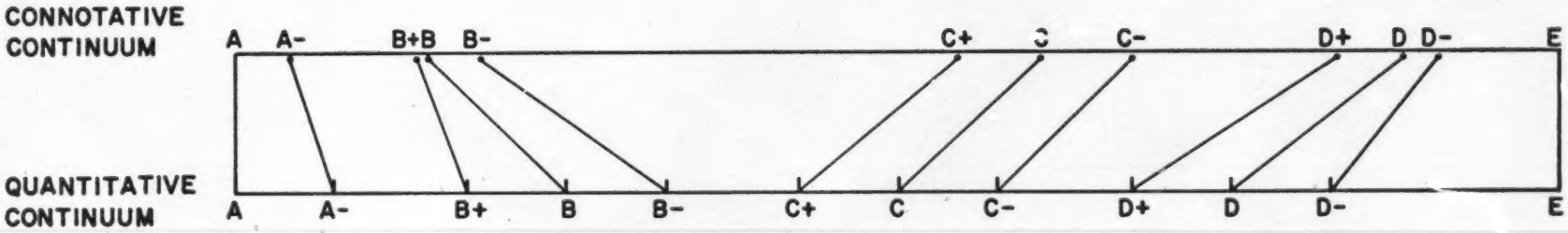

FIGURE I. DISTANCES BETWEEN GRADES ON THE CONNOTATIVE CONTINUUM AND THE QUANTITATIVE CONTINUUM